V2.0 Last updated January 27, 2021

HCMI Searchable Catalog User Guide

What is the purpose of this document?

The purpose of this document is to provide users with a resource to effectively navigate the Human Cancer Models Initiative (HCMI) Searchable Catalog.

What is HCMI?

The Human Cancer Models Initiative (HCMI) is a collaboration between the US National Cancer Institute (NCI) – part of the National Institutes of Health (NIH), Cancer Research UK (CRUK), Wellcome Sanger Institute (WSI), and foundation Hubrecht Organoid Technology (HUB). The goal of HCMI is to create up to 1,000 next-generation cancer models from patient tumors that are clinically and molecularly characterized. For more information about the HCMI program, please visit: https://ocg.cancer.gov/programs/HCMI.

What is the HCMI Searchable Catalog?

The HCMI Searchable Catalog allows users to browse and identify potential next-generation cancer models generated by HCMI for use in research. Links to available associated molecular characterization data, clinical and biospecimen data at the National Cancer Institute's (NCI) Genomic Data Commons (GDC), the European Genome-phenome Archive (EGA), and the 3rd party HCMI Model Distributor are available on each model page as the information is processed and validated.

HCMI Searchable Catalog URL: https://hcmi-searchable-catalog.nci.nih.gov/

Supported browsers include Chrome, Firefox, and Edge.

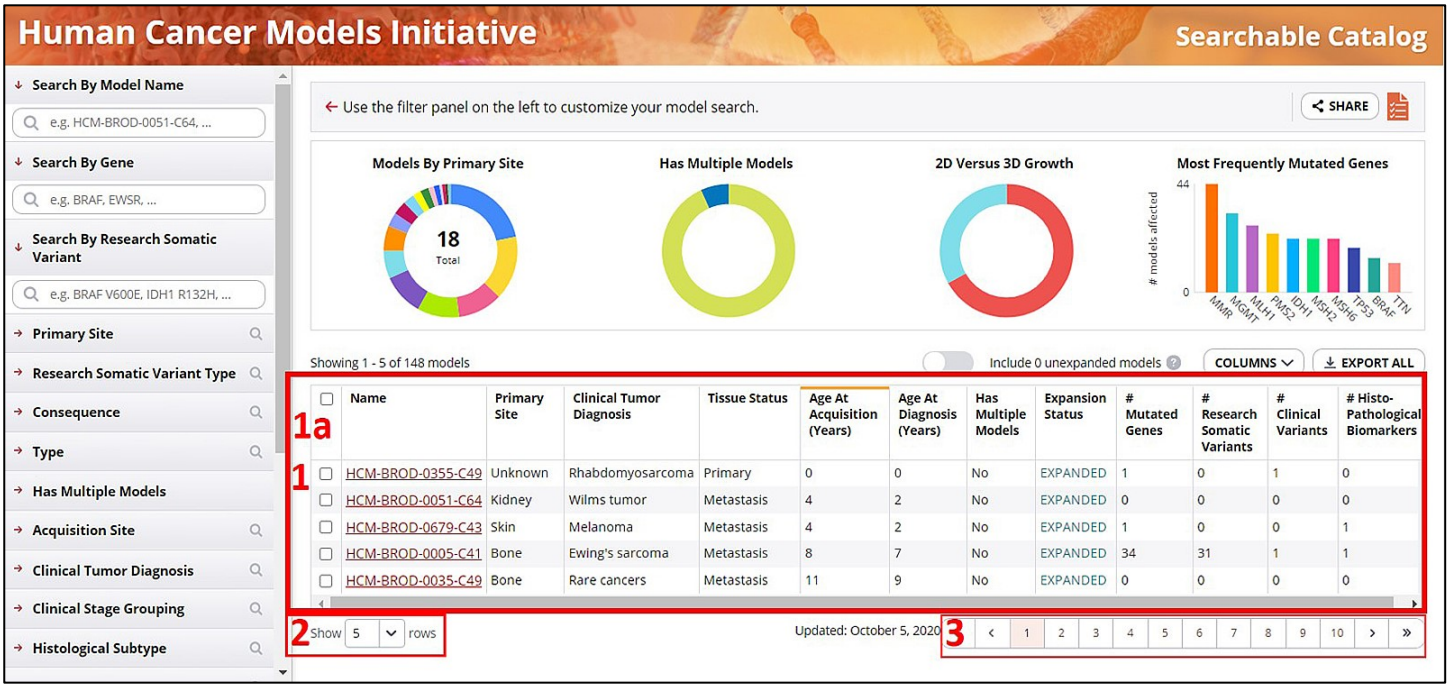
Table of Contents

V2.0 Last updated January 27, 2021

Navigating the HCMI Searchable Catalog Landing Page

What is displayed on the landing page?

The landing page features a dynamic view, enabling users to filter the models by several elements, and easily download the saved results.



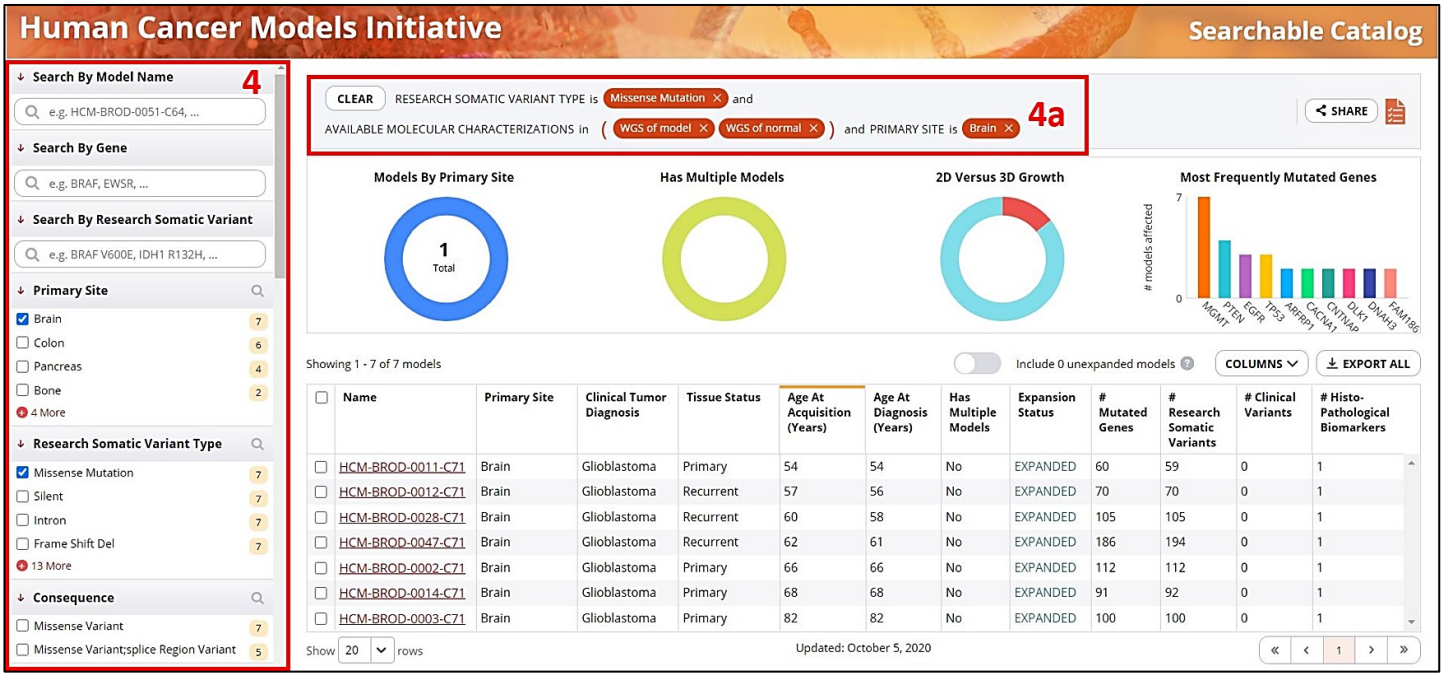
Box (**1**) shows the main viewing pane of the HCMI models. Users can sort the models in ascending or descending alphabetical or numerical order by clicking on the column headers such as “NAME”, “PRIMARY SITE”, etc.

Box (**2**) shows the number of models that users may choose to view.

Box (**3**) shows additional pages, if any, of available models.

Return to Table of Contents

Querying the Searchable Data Elements

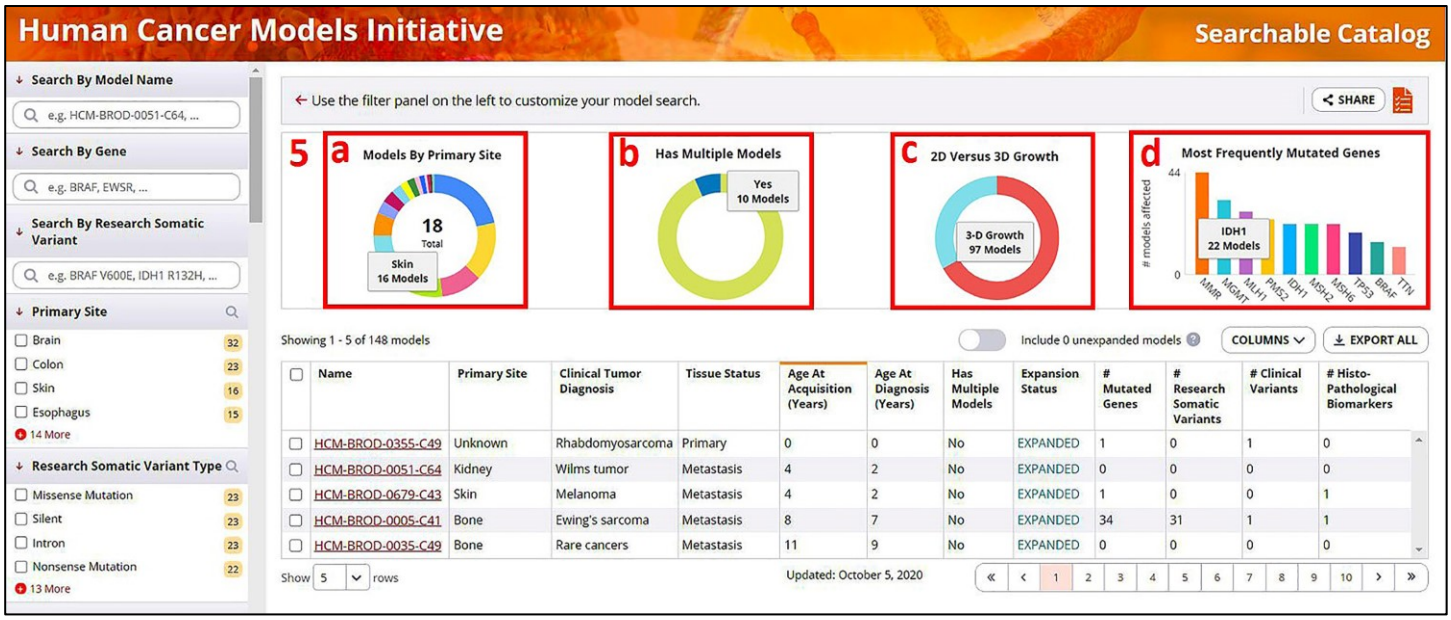


How do I filter data elements to query the available models?

Users may either search the models by typing in the model name, gene name or the research somatic variant type. Users may also filter the models using the data types displayed in Box (**4**) by checking or unchecking boxes to select or deselect data elements of interest. The list of models displayed in the main viewing pane and graphs will dynamically change as data elements in box (**4**) are selected or deselected.

Box (**4a**) at the top of the viewing pane shows the selected filtered data elements. Users may deselect a searchable data element by clicking the 'X' next to the name of the filtered element. To reset all selections, utilize the "CLEAR" button at the top of the pane.

Return to Table of Contents



What do the graphs on the main viewing page indicate?

Users may also filter the models by clicking on various colors within the interactive graphs shown in (**5**). Hovering over different colors within a graph will reveal relevant information.

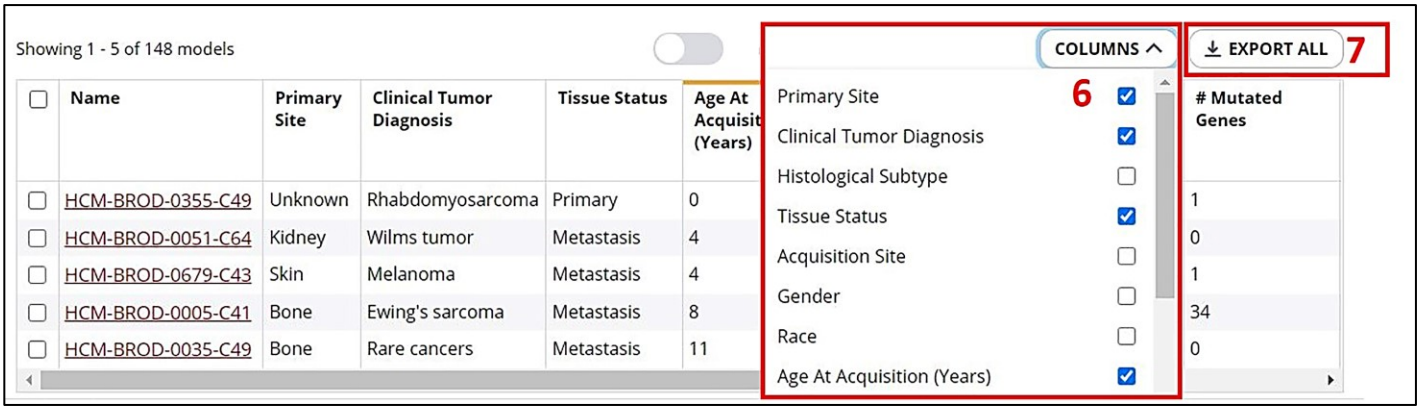
Box (**5a**) shows that users can click on different colors within the circle graph and filter the models by primary sites of the available models. The example shown here highlights “Skin” as a primary site.

Box (**5b**) shows that users can select the models by whether the cases have multiple models derived from independent tumors from the same patient (e.g. primary and metastatic, primary and pre-malignant tissues, etc.) or not. The example shown here has the mouse hovered over the “blue” indicating the number of models that have multiple models.

Box (**5c**) shows that users can query the models by the growth type: 2-D or 3-D. The example shown here features 3-D models.

Box (**5d**) shows that users can query the models within the main viewing table by selecting the most frequently mutated gene(s). Users may select more than one gene at a time. The example shown here indicates the number of models with IDH1 mutation information.

Return to Table of Contents

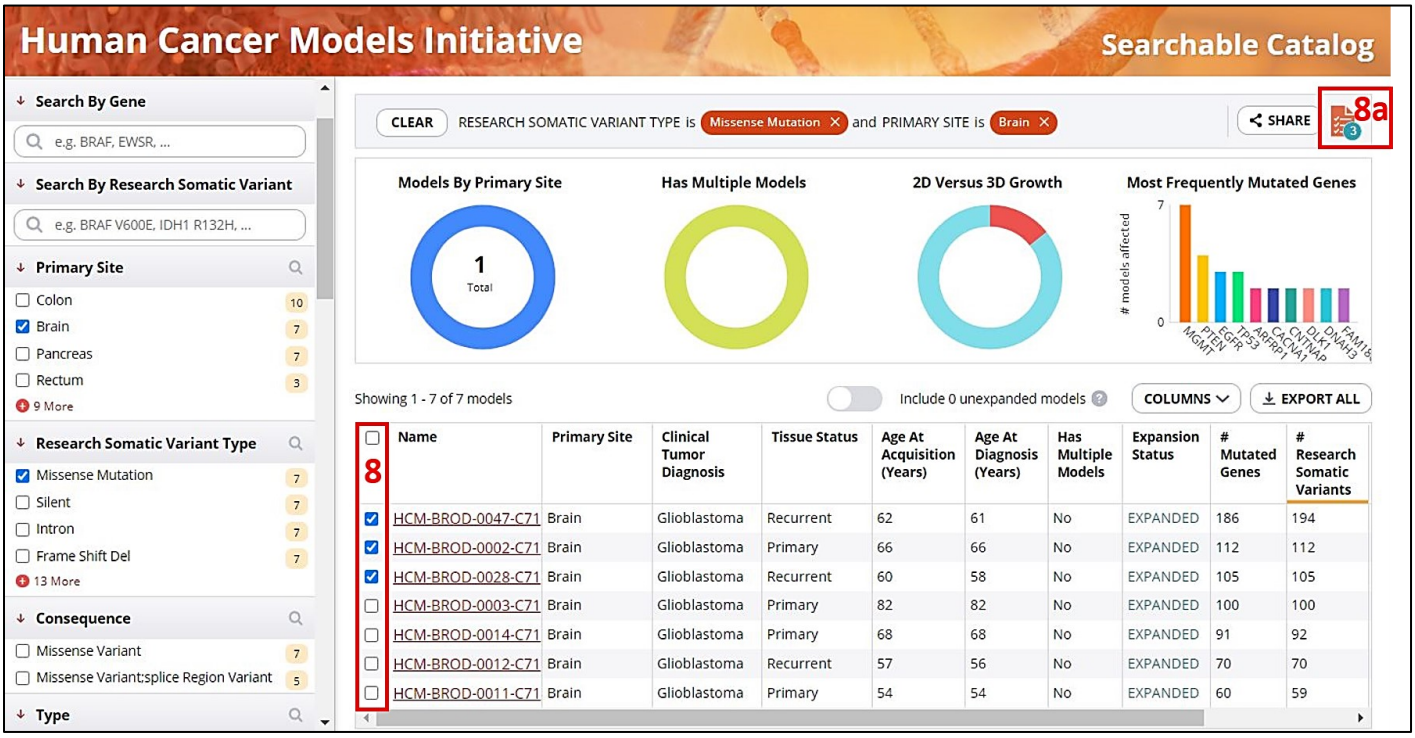


How can I download information regarding selected models?

The default columns shown in the main viewing pane can be customized based on information of interest. Users may select the data elements under the "COLUMNS" button shown in Box (**6**) to view in the main viewing pane.

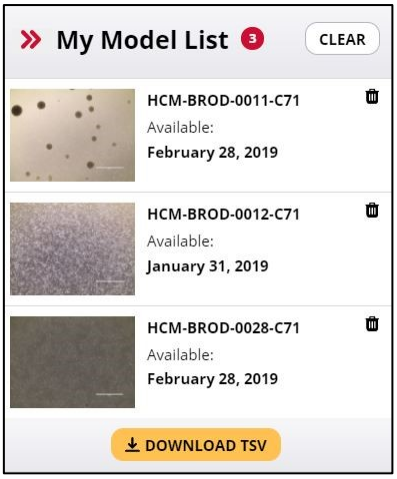
Once users have selected the data elements to be shown in the viewing pane, these data can be exported using the "EXPORT ALL" function in box (**7**) and saved as a .tsv file. To view the .tsv file in a tabular format, users may open the saved file in Excel or similar program.

Return to Table of Contents



How can I save models from multiple searches for later use?

In box (**8**), users may select the models of interest by clicking on the checkboxes next to the model name. In the example shown, the "Primary Site" of "Brain" is selected. Users may further sort and select models that meet certain characteristics, for example, "# of Research Somatic Variants". Selecting models adds them to the "My Model List" (box **8a**), which functions as a 'shopping list'. This feature may be helpful to users if multiple searches are conducted within a session. For example, if the user is also interested in models from the "Primary Site" of "Pancreas", an additional search may be conducted, and further models may be added to "My Model List".



Clicking on the icon in box **(8a)** allows users to view the models that have been saved to "My Model List". Users may download the list of saved models by clicking the DOWNLOAD TSV icon, individually delete the selected models by clicking the corresponding trash symbol, , or clear all selected models from the list by clicking "CLEAR".

Return to Table of Contents

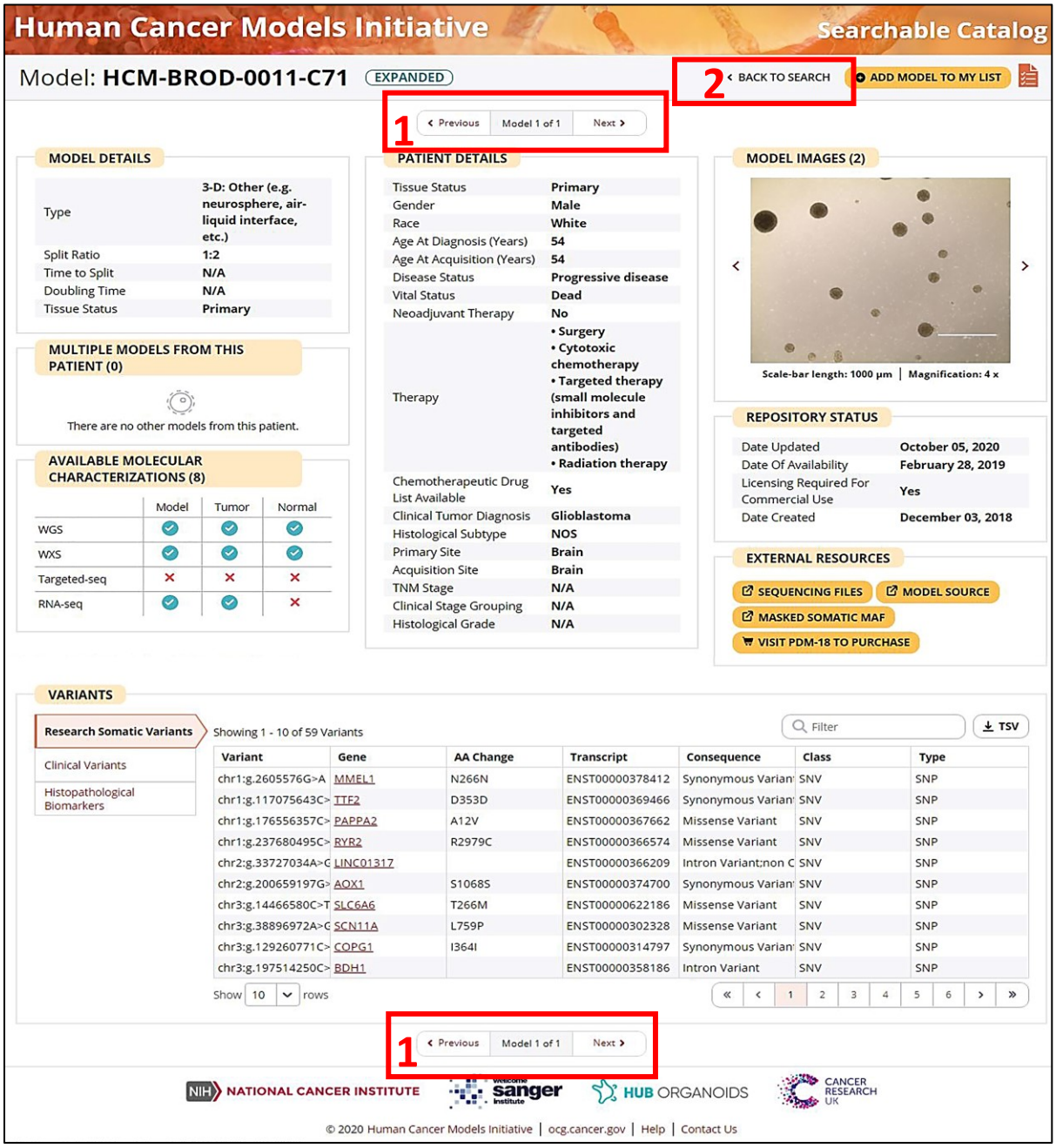
V2.0 Last updated January 27, 2021

Navigating Individual HCMI Model Pages

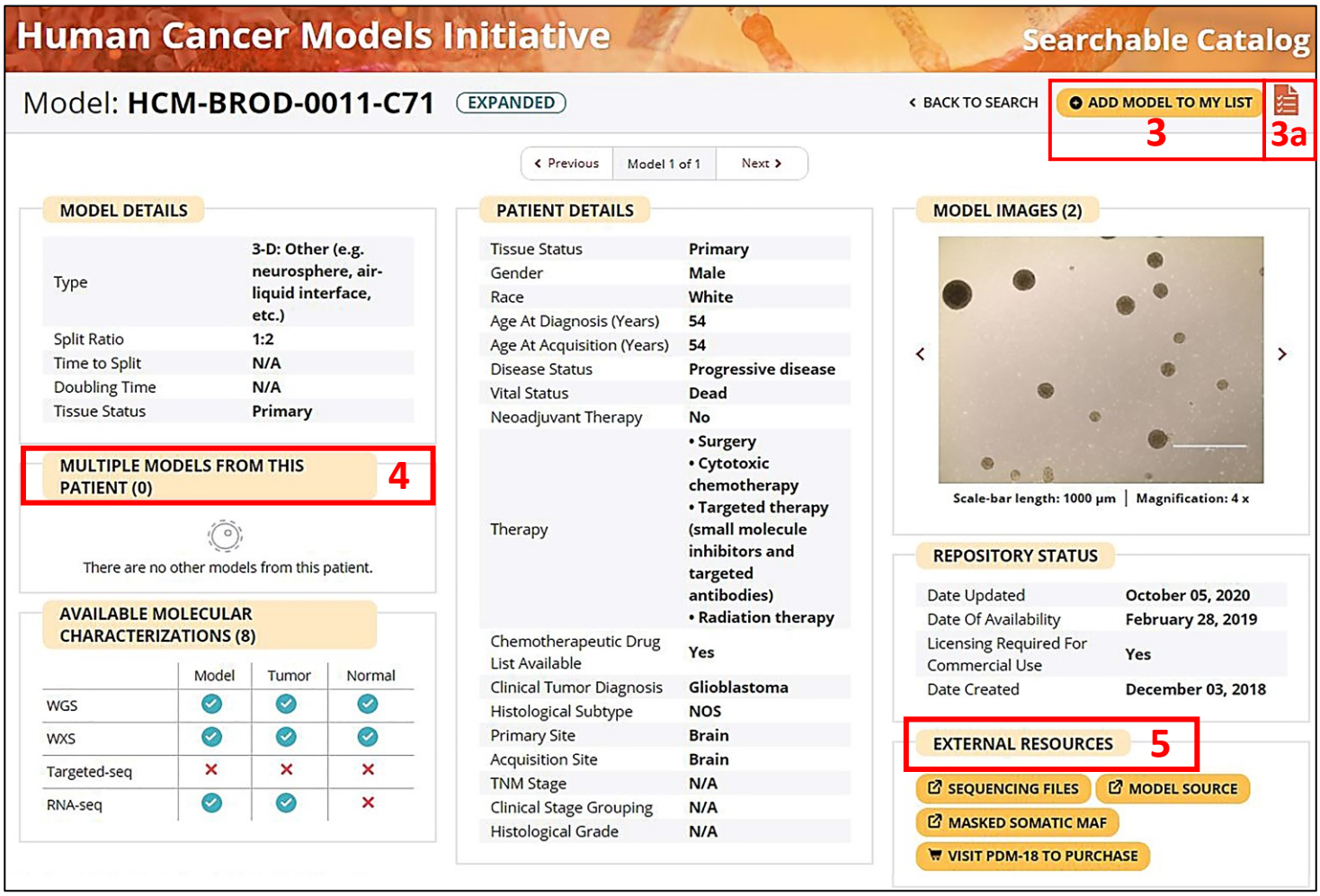
How do I navigate from one individual model page to the next?

To view an individual model page, users may select one of the models listed within the main viewing table by clicking on the model name of interest. In box (**1**), users may navigate to the previous or next individual model pages within the filtered list by clicking on the left/right arrows at the top or bottom of the page.

To return to the search results, users may select the “Back to Search” button located at the top of the page (box **2**).



Viewing Individual Model Data



How do I add models to "My Models List" from the individual model pages?

Individual models can be added to the "My Model List" by clicking ADD MODEL TO MY LIST (box **3**). To view and download the saved models, click the "My Model List" icon (box **3a**).

What information are available on individual model pages?

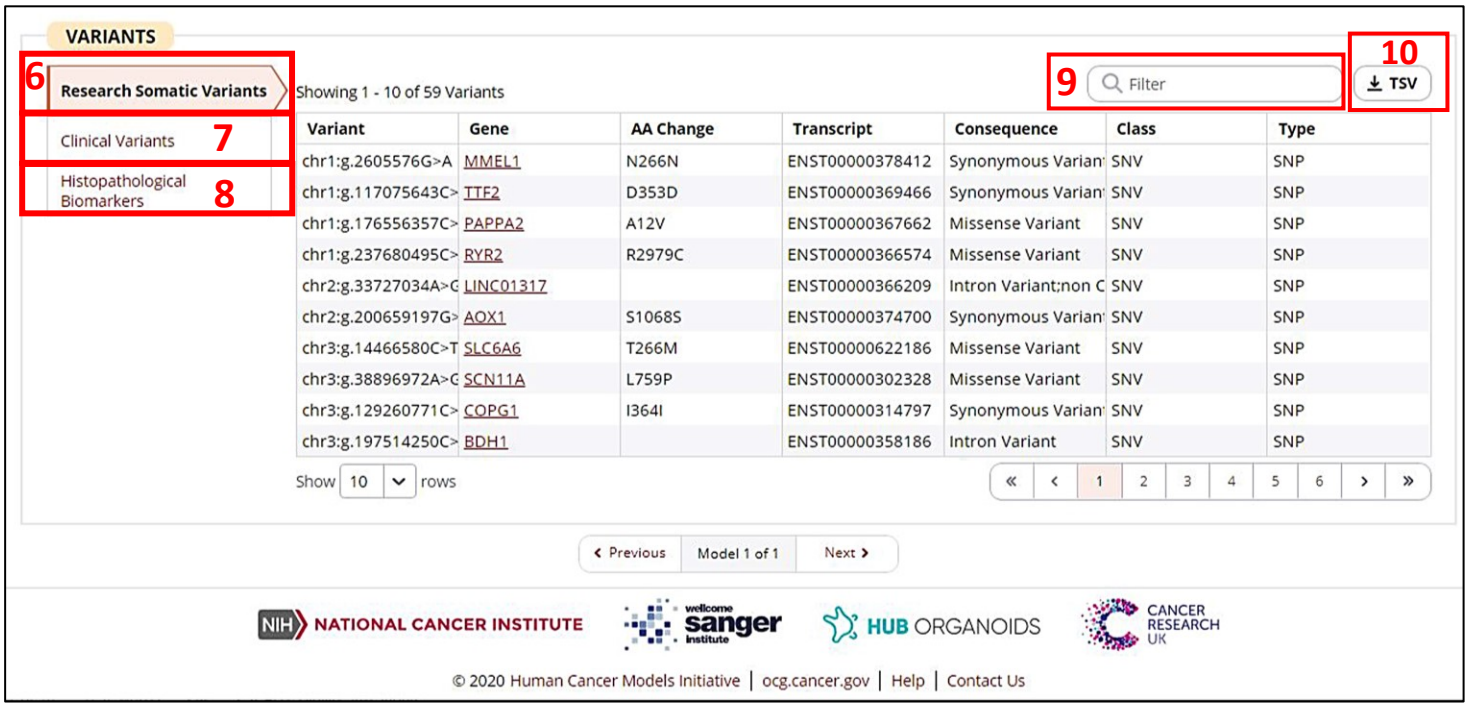
On each individual model page, all available data elements such as "MODEL DETAILS", "PATIENT DETAILS", and "MODEL IMAGES" are described.

What information do additional categories include?

"MULTIPLE MODELS FROM THIS PATIENT" (box **4**) indicates whether there are other models derived from independent tumors from the same patient (e.g. primary and metastasis, primary and pre-malignant, primary and recurrent, etc.).

"EXTERNAL RESOURCES" (box **5**) contains links to available sequencing data at the GDC, corresponding model page at the GDC, models' masked somatic MAF data page, and the model distributor's page.

Return to Table of Contents



What information does the "VARIANTS" section contain?

The "VARIANTS" section contains available "Research Somatic Variants", "Clinical Variants", and "Histopathological Biomarkers" data tabs. The data columns within each tab can be sorted either in ascending or descending order by clicking on the column header of interest.

"Research Somatic Variants" tab (box **6**) shows available open-access masked somatic MAF variant data at the GDC. As part of GDC's harmonization process, potential germline mutations are filtered from the variants identified by DNA sequencing of the model and normal tissues. These highly-filtered lists of somatic mutations without the germline variants are called "masked somatic mutations". The masked somatic mutations generated at GDC for each model are shown as "Research Somatic Variants" on the Catalog. Catalog users may search the available models for gene mutations of interest. If omission of true-positive somatic mutations is a concern, it is recommended that users access the controlled-access MAF files housed at the GDC. Access to controlled-access data is granted through dbGaP. Visit the "Accessing HCMI Data" page for more information.

The "Clinical Variants" tab (box **7**) shows available clinical variants reported from clinical sequencing of the tumor collected from the clinical record.

The "Histopathological Biomarkers" tab (box **8**) shows the results of reported clinical histopathological biomarkers collected from the clinical record.

Can the "VARIANTS" data be queried or downloaded?

In box (**9**), users may filter the "VARIANTS" information on each tab by entering querying text (e.g. MSH6, TP53, etc.). Users may download the variant information by clicking the TSV icon (box **10**).

Return to Table of Contents

Additional Help, Feedback, and Bug Reporting

Who do I contact if I have questions?

For questions, users may visit the HCMI FAQs page by clicking the "Help" button at the bottom of the webpage.

How do I report a bug?

Users may provide feedback or report bugs directly to the Office of Cancer Genomics by emailing: ocg@mail.nih.gov or by clicking "Contact Us" at the bottom of the Catalog webpage.

Return to Table of Contents